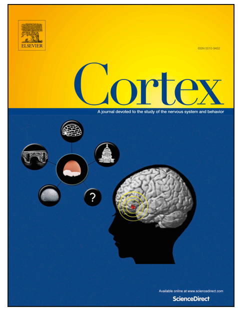
Accepted Manuscript



Sounds from seeing silent motion: Who hears them, and what looks loudest?

Christopher J. Fassnidge, Elliot D. Freeman

PII: S0010-9452(18)30074-1

DOI: 10.1016/j.cortex.2018.02.019

Reference: CORTEX 2265

To appear in: *Cortex*

Received Date: 24 October 2017

Revised Date: 16 January 2018

Accepted Date: 23 February 2018

Please cite this article as: Fassnidge CJ, Freeman ED, Sounds from seeing silent motion: Who hears them, and what looks loudest?, *CORTEX* (2018), doi: 10.1016/j.cortex.2018.02.019.

Sounds from seeing silent motion: Who hears them, and what looks loudest?

Christopher J Fassnidge, Elliot D Freeman

Cognitive Neuroscience Research Unit, City University of London, UK

Corresponding author:

Dr Elliot D Freeman
City, University of London
Northampton Square
London EC1V 0HB
United Kingdom
email: elliot.freeman@city.ac.uk

Keywords:

Synaesthesia; individual differences; Audiovisual perception; Synaesthesia

Significance

Some people hear sounds when watching things moving. We report a large-scale survey of this barely-known ‘visually-evoked auditory response’ (vEAR).

We have analysed what individual traits predict vEAR, and which visual stimuli evoke the strongest sounds. Results inform ongoing debates about the neural underpinnings of synaesthesia.

Even videos showing meaningless motion evoke sounds. This implicates relatively direct pathways that bypass complex scene interpret.

vEAR is also associated with auditory-evoked flashes, musical imagery and tinnitus-like experiences. General cortical excitability might explain these diverse phenomena better than specific anatomical abnormalities.

vEAR seems highly prevalent. This makes it easier to study than other rarer forms of sensory crosstalk, vEAR provides a convenient new platform for investigating the neural bases of normal and anomalous multi-sensory perception.

Abstract

Some people hear what they see: car indicator lights, flashing neon shop signs, and people's movements as they walk may all trigger an auditory sensation, which we call the visual-evoked auditory response (vEAR or 'visual ear'). We have conducted the first large-scale online survey (N>4000) of this little-known phenomenon. We analysed the prevalence of vEAR, what induces it, and what other traits are associated with it.

We asked respondents if they had previously experienced vEAR. Participants then rated silent videos for vividness of evoked auditory sensations, and answered additional questions.

Prevalence appeared higher relative to other typical synaesthesias. Prior awareness and video ratings were associated with greater frequency of other synaesthesias, including flashes evoked by sounds, and musical imagery. Higher-rated videos often depicted meaningful events that predicted sounds (e.g. collisions). However, ratings were also driven by the low-level 'motion energy' of non-predictive flashing or moving patterns, specifically in respondents who had previous awareness of vEAR.

Our motion energy analysis suggests that signals from visual motion processing may affect audition relatively directly, without requiring higher-level interpretative processes. While some popular explanations of synaesthesia assume rare and specific patterns of brain hyper-connectivity, the apparently high prevalence of vEAR, and its broad association with other synaesthesias and traits, are consistent with a common dependence on normal variations in physiological mechanisms of disinhibition or excitability of sensory brain areas and their functional connectivity, rather than just on specific patterns of hyper-connectivity. The prevalence of vEAR makes it easier to test such hypotheses further, and makes the results more relevant to understanding not only synaesthetic anomalies but also normal perception.

Introduction

It is well known that what we see can influence what we hear. For example the sight of a person's lip movements can enhance speech comprehension or even change our interpretation of the speech sounds (McGurk & MacDonald, 1976; Sumby & Pollack, 1954), while the movements of a musician can dominate the perceived quality of their performance even among expert listeners (Tsay, 2013). However it is much less appreciated that vision not only modulates perception of concurrent auditory stimuli, but sometimes can also induce the perception of new auditory sensations (Saenz & Koch, 2008). For example, some individuals claim they can 'hear' flashing car indicator lights or shop displays, or people's movements while walking or speaking. Over the last decade since Saenz & Koch's (2008) short report, there has been remarkably little further research on this intriguing phenomenon (Fassnidge, Cecconi Marcotti, & Freeman, 2017; Rothen, Bartl, Franklin, & Ward, 2017), which we call the 'visually-evoked auditory response' (vEAR, also known as 'hearing motion synaesthesia'). Many questions therefore remain to be answered. For example, it is currently unknown whether vEAR represents a form of high-level cognitive association or imagery, or a genuine form of synaesthesia-like sensory crosstalk. It is also unknown how prevalent it is in a large population, what traits characterise people who experience vEAR, and whether normal or abnormal brain mechanisms might be responsible for it. We address these questions here using the results of a large-scale internet survey. We have assessed the prevalence of vEAR, and for the first time analysed the kinds of visual motion stimuli that evoke high ratings of auditory sensations, and the individual traits that predict high ratings and vEAR susceptibility.

Our study is relevant to two ongoing debates about the neural underpinnings of synaesthesia (Hubbard & Ramachandran, 2005). One debate concerns the neuro-architecture that results in sensory cross-talk: whether synaesthesia is mediated by feedback from high-level semantic representations, or whether it involves relatively more direct cross-wiring between sensory modalities. For example, some have argued in favour of direct anatomical connections between brain areas (Ramachandran & Hubbard, 2001), while other evidence suggests that this phenomenon requires prior semantic interpretation of the inducing stimulus (Mattingley, Rich, Yelland, & Bradshaw, 2001; Myles, Dixon, Smilek, & Merikle, 2003; Smilek, Dixon, Cudahy, & Merikle, 2001). A second debate concerns the neurophysiological causes of synaesthesia: it might result from unusual anatomical patterns of connectivity (Bargary & Mitchell, 2008; Baron-Cohen, 1996; Hubbard & Ramachandran, 2005; Tomson et al., 2011), or from more systemic physiological variables which disinhibit normally-occurring connections between sensory areas or render the areas themselves more excitable (Grossenbacher & Lovelace, 2001; Neufeld et al., 2012).

Concerning neuro-architecture, the role of semantic representations can be ambiguous in some synaesthesias which involve inducers that are relatively high-level or cultural in origin, such as letters or words evoking colours (Bor, Rothen, Schwartzman, Clayton, & Seth, 2014; Witthoft, Winawer, & Eagleman, 2015). There is potentially less ambiguity where synaesthesia involves associations between more basic sensory dimensions such as sound and colour, or in this case between visual movement and sound. However in vEAR there may still be two routes to inducing a sound: one which depends on relatively direct crosstalk between areas processing low-level visual motion cues and audition, and a higher-level route that depends on prior semantic analysis of the visual scene and predictions about whether the depicted events are likely to produce sounds, associated for example with friction, collisions or explosions. Fortunately, in vEAR it is straightforward to quantify how much low-level 'motion energy' is present in the image sequence as the patterns of light changes over space and time, for example using a computational approach which models the spatiotemporal sensitivity of cells in early visual cortex to moving patterns (Adelson & Bergen, 1985). The contribution of this motion energy to the perception of vEAR may be measured independently from the higher-level semantic content of the images.

Concerning the second debate about the neurophysiological causes of synaesthesia, the assumption that synaesthesias depend on abnormal neural connectivity is supported by evidence that typical synaesthesias are both fairly rare and highly idiosyncratic (Simner et al., 2006; Ward, 2013). However another reason for the apparent rarity of synaesthesia might be that it is unusual to regularly encounter in nature the specific combinations inducers and concurrents that are typically associated in synaesthesia (Fassnidge et al., 2017). For example, consistent pairings between specific letters and colours do not occur in the natural environment, although grapheme-colour associations may be reinforced following repeated exposure in childhood to coloured fridge magnets (Bor et al., 2014; Witthoft et al., 2015). On this statistical view, one might expect higher prevalence of vEAR than other synaesthesias because vision and audition are naturally highly correlated with each other, for example every time our footstep hits the ground or we watch a person speaking. Indeed, we previously found that 22% of our lab noticed faint sounds evoked by silent 'Morse-code' flashes (Fassnidge et al., 2017), however this was a small sample (N=37). If vEAR were found to have greater prevalence than other typical synaesthesias, this would suggest that some synaesthesia-like phenomena could occur via relatively normal rather than rarely occurring patterns of neural connectivity.

Despite the individual idiosyncrasy of synaesthetic associations, there is evidence that people with one kind of synaesthesia are more likely to report others (Barnett et al., 2008; Rothen et al., 2013; Sagiv, Simner, Collins, Butterworth, & Ward, 2006), as well as evidence of distinct personality profiles (Banissy et al., 2013; Rouw & Scholte, 2016), comorbidities for example with schizotypy and autism spectrum disorders (Banissy et al., 2012; Baron-Cohen et al., 2013; Ward et al., 2017), and other generalised benefits in sensory acuity (Banissy, Walsh, & Ward, 2009). Such broad patterns of association would be more supportive of the notion that there are systemic variables, possibly of a genetic origin, governing the expression of synaesthetic phenomena and its associated traits (Barnett et al., 2008; Carmichael & Simner, 2013). Such variables might broadly affect development of connectivity, and/or impact on cortical excitability, which has independently been linked to diverse crossmodal and synaesthetic phenomena (Bolognini, Senna, Maravita, Pascual-Leone, & Merabet, 2010; Schroeder, Lakatos, Kajikawa, Partan, & Puce, 2008; Terhune, Tai, Cowey, Popescu, & Cohen Kadosh, 2011). To further explore such associations, we aimed to test here whether vEAR correlates with other synaesthesias and sensory phenomena.

Our empirical method involved eliciting ratings of the vividness of auditory sensations evoked by silent videos. We assessed the extent to which the semantic versus low-level characteristics of the videos (as indexed by a measure of motion energy) each contributed to ratings of the vividness of any evoked auditory sensations. We asked about prior experience of vEAR in order to assess prevalence. We also briefly asked about other kinds of synaesthesia, involuntary musical imagery (Kumar et al., 2014), flashes evoked by sudden sounds (Jacobs, Karpik, Bozian, & Gøthgen, 1981; Lessell & Cohen, 1979), and tinnitus (Kaltenbach, 2011). These phenomena have all been independently associated with raised cortical excitation or disinhibition.

We reasoned that if reports of vEAR were based only on high-level semantically-mediated associations, then higher ratings should be given only to videos depicting events which are naturally associated with sounds, such as fireworks exploding, a person shouting, or collisions. This form of high-level association might be hard to distinguish from cognitively-mediated imagery. Alternatively, if vEAR depends on audiovisual connections that bypass high-level scene interpretation, then ratings might also depend on the amount of low-level motion energy in the videos (Adelson & Bergen, 1985). Silent videos with higher motion energy, such as flashing neon shop signs, twinkling LEDs, or abstract swirling dot patterns, might attract higher ratings, even if the depicted events are not naturally accompanied by sound.

If vEAR is prevalent, this would challenge the notion that synaesthetic associations only reflect rare and abnormal hyperconnectivity between specific brain areas, while supporting more an alternative account in which connections between vision and audition are already naturally rich in many individuals, even if not always fully functional. Furthermore, if vEAR is associated with other diverse traits, this would be harder to explain with the assumption of idiosyncratic neural connections between specific brain areas, but instead support the role of more general dimension of individual variability, such as reduction of sensory inhibition or increased excitability in audiovisual areas.

Methods

Participants

Procedures were approved by the Psychology ethics review board at City University of London. This survey was publicised via on-line news reports from the popular press (e.g. Devlin, 2017) about our previous publication on the topic of vEAR (Fassnidge et al., 2017). Participants were recruited to the survey after following a hyperlink within the news reports, advertised with text such as '*Do you experience "hearing motion"? Take the test here*: tinyURL.com/vEARsurvey'. No compensation was offered for participating. Location data from respondents showed world-wide participation, although dominated by North America, the United Kingdom, and Western Europe. A total of 33,504 individuals arrived at the first informed consent page of the survey, however only 4128 completed the whole survey. 3212 of the non-completers quit the survey before answering the first question, and on average 26118 attempted no more than 3 questions. After the initial release of the survey, we appended some further trait-related questions, which 1058 participants completed. Demographic data are reported in **Error! Reference source not found.**. To assess possible self-selection bias, a naïve sample of 132 paid respondents were later recruited by advertisement from a participant panel (Prolific.ac) and the local Psychology department. The recruitment advertisement referred briefly to 'synaesthetic abilities' but made no mention of vEAR. 126 out of 132 completed the full survey.

Materials

The survey was administered on-line using Qualtrics, and may be viewed at http://tinyurl.com/vEARsurvey (see also Figure 1). Our stimuli consisted of 24 royalty-free high-definition video clips downloaded from www.videoblocks.com. We selected a variety of videos depicting meaningful versus abstract subjects engaged in slow, fast, smooth or sudden movements. We trimmed their duration to 5 seconds, and cropped them to 640x360 pixels. The actual visual angle of the stimuli was not possible to control, as they were rendered on a variety of displays. Examples included a ballet dancer performing a pirouette, bouncing balls, a hammer hitting a nail, a person screaming, and flashing neon displays, as well as more abstract animations such as swirling patterns of dots.

The survey included additional multiple-choice questions probing demographics, experience of vEAR, and other traits. All participants were asked about their age, gender, and whether they considered themselves to be synaesthetes, with some typical examples of synaesthesia provided. Non-completing participants typically quit the study before answering these. A subset of 1058 participants recruited from the same source (plus the 126 paid participants), were asked an additional set of questions: *"Do you suffer from tinnitus (ringing in the ears)?*", *"When in the dark or falling asleep, do you ever see flashes of light triggered by sudden sounds?*" (hypnagogic auditory-evoked phosphenes), "*Do you ever hear music in your head?",* and *"In everyday life are you ever aware of hearing sounds when you see flashing lights or movement? (e.g. shop displays, car indicators, or people walking?)"*. We also asked whether participants experience synaesthetic associations, giving some canonical examples: "'*Synaesthesia is a rare condition where sensation in one sense can cause you to experience sensation in another sense. Examples might include seeing colours when you hear music, always seeing particular letters and numbers in specific colours, or experiencing tastes/smells when you hear or read particular words. Do you consider yourself to be a be a synaesthete?"*. In the interests of limiting the length of the survey, we did not undertake a detailed survey of specific types of synaesthesia, or an assessment of other traits using detailed standardised measures.

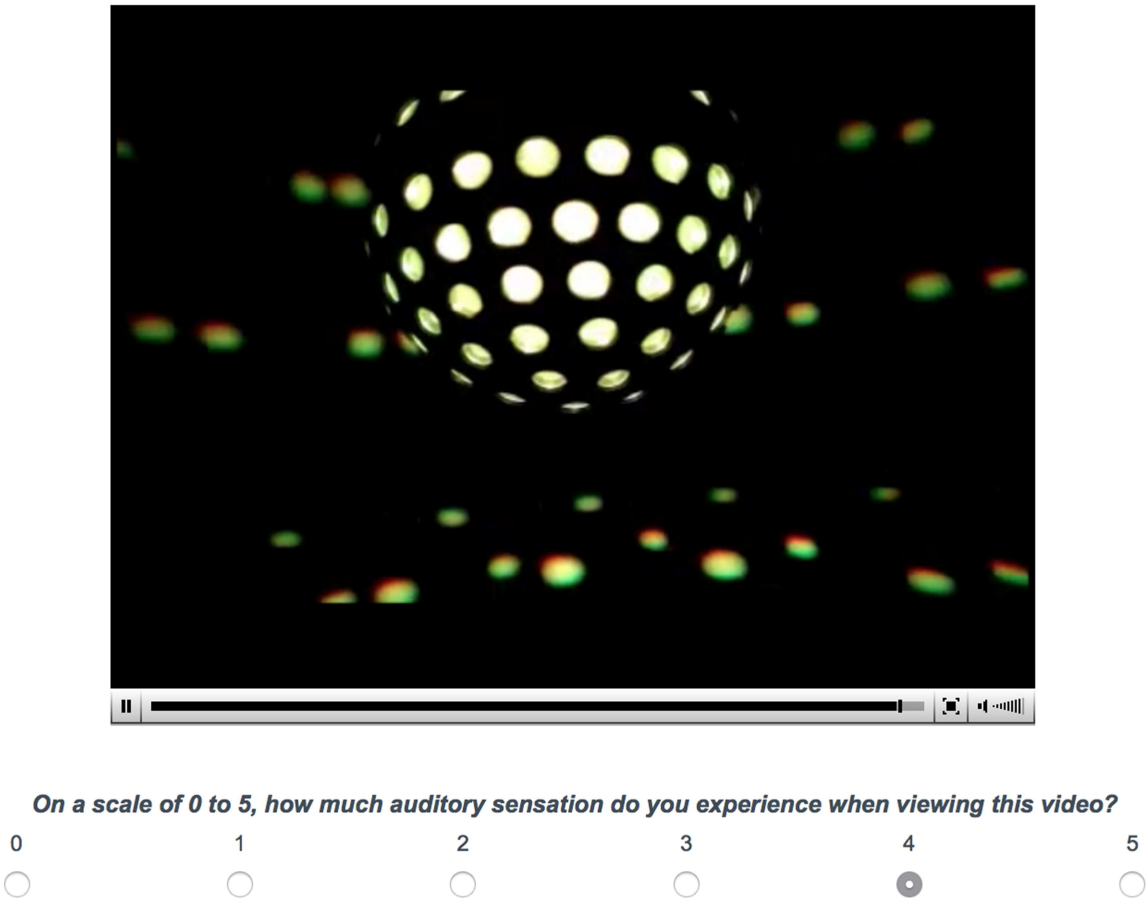


Figure 1 Example page from on-line survey showing movie and response options.

<u>Procedure</u>

Following an informed consent page, participants were then shown a briefing page explaining the nature of vEAR and how it might differ from normal hearing. The text, transcribed below, compiles descriptions of the phenomenon from the original research report (Saenz & Koch, 2008), from the authors' informal interviews, and the everyday experience of one of the authors (EDF):

> *The sound may be experienced within your head rather than in the outside environment. You may experience it as if you are vividly imagining the sound, or it may sound like a ringing in your ears, or it might resemble the experience of 'hearing' phrases of a popular song in your mind's ear, or the voices of people on television when watched with the volume off. Alternatively it may be an abstract experience, but closer to being an auditory experience than a visual experience. Some people describe it as imaginary white noise. What is important is that the auditory sensation occurs in time with visual change over time, caused by motion or sudden flashes. It is typically*

> *involuntary (i.e. it happens automatically rather than as a result of conscious effort) and it happens consistently. Have you previously been aware of experiencing this type of auditory sensation when viewing visual movement? [No / Not sure / Yes]*

Participants were then asked to watch each video and then asked: “On a scale from 0 to 5, how much auditory sensation do you experience when viewing this video?”. Instructions were to use rating 0 for “no auditory sensation at all”, and 5 for “very vivid and definite auditory sensation”. Participants were instructed that they could repeat the video playback as many times as required before rating them. They were also informed that the sensations may be very faint and they might have to listen carefully. They were also asked to work in a quiet environment. Videos appeared in the centre of the screen, with order randomised between participants. Once a rating had been made, participants could continue to the next video. Following the videos, there were further questions (see Stimuli).

Results

Prevalence

All respondents answered the first question about whether they were previously aware of experiencing auditory sensations accompanying visual movement. Of the 26,118 who did not complete the survey, 16% answered 'Yes'. Of the remaining 4128 respondents who did complete the survey the proportion of 'previously aware' participants was 21% See **Error! Reference source not found.** for more detailed results. This proportion was significantly higher than in the non-completing respondents [χ^2 = 59.78, p<.0001], and might reflect self-selection by participants who were motivated to experience vEAR or curious to find out more about their condition. 12% of completing participants reported identifying as synaesthetes, although this should be considered with caution given that prevalence estimates based on informal questions rather than standardised assessments are often greatly inflated (Simner et al., 2006).

	Sampling		Awareness of vEAR			
Main sample	**Video rating only**	**Video rating + questionnaire**	**No**	**Not Sure**	**Aware**	**All**
Incomplete			10740	11210	4168	26118
%			41%	43%	16%	
Complete	3070	1058	1566	1698	856	4128
%			38%	41%	21%	
Age Mean (SD)	38.1 (13.9)	35.8 (13.3)	40.4 (14.3)	35.9 (13.2)	35.5 (13.0)	37.5 (13.8)
Female	43%	51%	40%	48%	48%	45%
Male	56%	47%	59%	51%	50%	54%
Naïve paid sample						
Incomplete			6	0	0	6
%			100%	0%	0%	
Complete			50	37	39	126
%			40%	29%	31%	
Age Mean (SD)			35.2 (11.9)	34.6 (13.3)	27.4 (7.2)	32.6 (11.6)
Female			56%	46%	49%	50%
Male			44%	54%	51%	49%

Table 1 Demographics.

To assess whether self-selection bias could have inflated prevalence estimates, we administered the same survey to an additional small sample of 126 naïve paid participants (mean age 32.6, SD 11.62, 64 female). These participants were motivated by monetary reward rather than interest in vEAR, and showed much less attrition than in the main sample, thus it seems reasonable to assume that this was a relatively unbiased sample. The proportion of these participants responding ‘Yes’ to the initial question about prior awareness of vEAR was actually slightly higher than in the main sample (‘Yes’, 31%; ‘Not sure’, 29%; ‘No’, 40%), even though these respondents were not informed at recruitment about the subject of the study. The proportion of this sample identifying as synaesthetes was 5%. The pattern of ratings and responses to the additional questions showed very similar trends compared to the large unpaid sample, with a similar pattern of statistically significant results, but we focus on the latter much larger group in the results reported below.

To avoid potential biases in interpreting the subjective question about previous awareness of vEAR, we adopted an alternative method to estimate prevalence, following a recent smaller-scale survey (Rothen et al., 2017) in which participants were asked whether or not they experienced sounds while watching each of 12 videos. The diagnostic criterion was that they had to answer ‘Yes’ to at least half of the videos. Using this method, Rothen et al (2017) estimated prevalence at 4.2%. We used a similar method, whereby respondents had to rate at least 12 of our 24 videos with values equal to or greater than a given criterion rating. In ascending order of criterion rating (1 to 5), prevalence estimates from our naïve paid sample are as follows: 74.60%, 45.24%, 22.22%, 8.73%, 1.59%. Our main sample showed very similar values. These prevalence estimates are higher than Rothen et al’s for all but the most conservative criterion.

Demographics and trait frequencies

There was a significant association of age with responses to the initial question about previous awareness of vEAR [$F(2, 4117) = 57.91$, $p<0.0001$, $\eta^2_{part} = 0.027$]. ‘Yes’ respondents were younger than ‘No’ respondents by about 5 years. Crosstabulation of these responses with gender showed higher proportion of ‘No’ respondents in males [$\chi^2(2)= 38.17$, $p<0.001$, Cramer’s V = 0.08, excluding 51 respondents who selected neither male nor female categories]. See Figure 1 for means and frequencies.

Respondents to the extended end-of-survey questionnaire (N=1058) who reported previous awareness of vEAR were significantly more likely to report experiencing auditory sensations from flashing lights or movements in everyday life [$\chi^2(4)$ = 245.73, V = 0.30], musical imagery [$\chi^2(8)$ = 93.79, V = 0.11], sound evoked flashes [$\chi^2(4)$ = 90.89, V = 0.17], synaesthetic associations [$\chi^2(4)$ = 75.06, V = 0.16, and tinnitus [$\chi^2(4)$ = 19.96, V = 0.08], all p<0.0001, Figure 2). A repeated analysis excluding 'not sure' responses showed a similar pattern of significant associations. We also obtained a very similar pattern of significant results (including the analyses described below) when including only respondents who responded consistently to the questions about previous awareness and about hearing flashes or movement in everyday life (accounting for 53.85%). Of this reduced sample of 567 respondents, 173 responded 'Yes' to both questions (31%), and 172 responded 'No'.

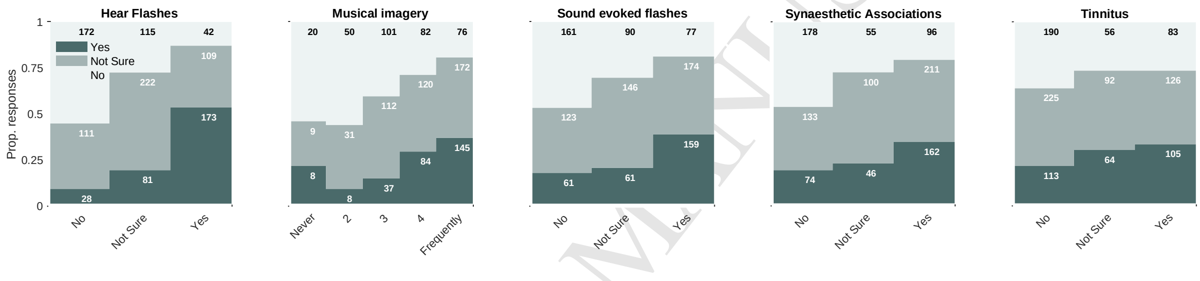


Figure 2 Proportion of 'previous awareness of vEAR' crosstabulated against traits probed at the end of the survey. Y-axis on left is common to all graphs. Different shades represent respondents with differing reported levels of previous awareness, with proportions displayed as stacked bars, for each category of response to the trait questions.

Ratings

Rating of visually-evoked auditory sensations had very high internal consistency (Cronbach's Alpha = 0.97). Average rating in the full completed sample was 1.46 (SD 1.18). Ratings were significantly higher in participants reporting previous awareness of hearing visual motion [t(4126) = 26.74, p<0.00001, Cohen's D = 1.027]. Rating were slightly higher in younger participants [F(5,4122) = 28.94, p<0.0001, η^2_{part} = 0.018], and in females [t(4075) = 8.16, p<0.00001, Cohen's D = 0.257].

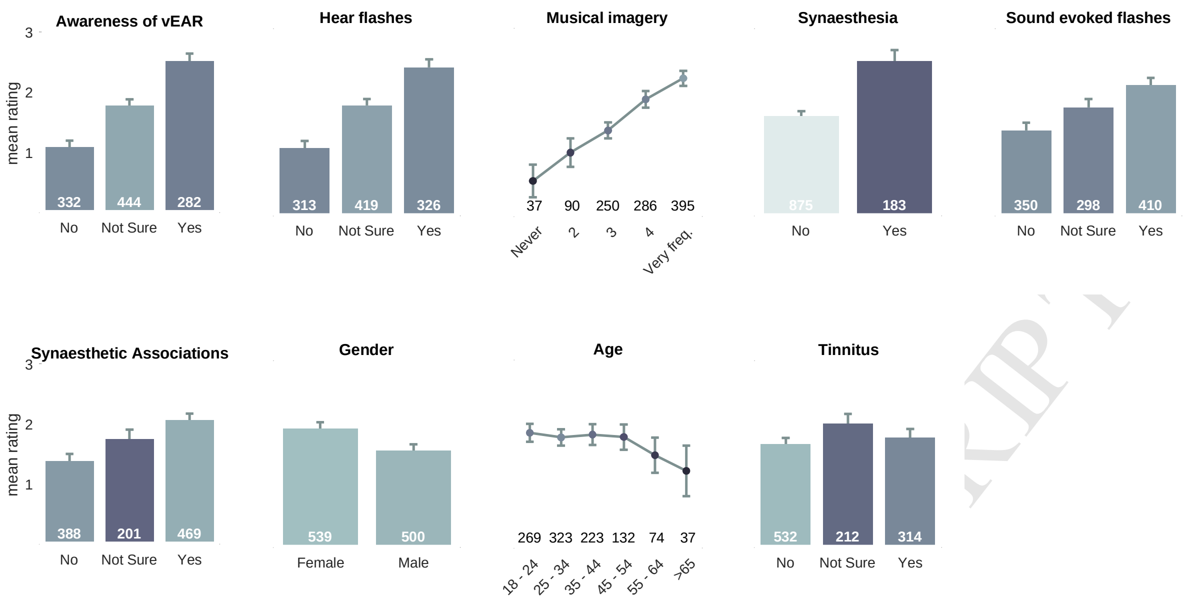


Figure 3 Mean video ratings broken down by responses to trait and demographic variables, with 95% confidence intervals (N=1058). Y-axis on left is common to all graphs. Variables displayed in order of effect size; tests for overall differences between means (F) are all significant [p<.05]. Shading of bars and numbers at base of graphs indicate numbers of respondents in each category.

Effects of traits on ratings are shown in **Error! Reference source not found.** for the 1058 respondents who completed our end-of-survey questionnaire. All main effects of trait were highly significant ($p<.0001$, except where indicated below), although none interacted significantly with previous awareness of vEAR. Previous awareness of vEAR was a determinant of higher ratings [$F(2,1055) = 132.96$, $\eta^2_{part} = 0.201$], as well as reports of hearing flashes and movements in everyday life [$F(2,1055) = 111.41$, $\eta^2_{part} = 0.174$]. In addition, ratings were lower with age [$F(5,1052) = 2.52$, $p<0.028$, $\eta^2_{part} = 0.012$] and higher in females [$F(1,1037) = 23.19$, $\eta^2_{part} = 0.022$], while ratings showed a clear positive association with frequency of experiencing musical imagery [$F(4,1053) = 43.72$, $\eta^2_{part} = 0.142$]. Significantly higher ratings (although with smaller effect sizes) were also made on average by participants who claimed to have a form of synaesthesia [17% of respondents, $F(1,1056) = 86.11$, $\eta^2_{part} = 0.075$], or to have experienced varieties of synaesthetic associations [44%; $F(2,1055) = 35.53$, $\eta^2_{part} = 0.063$]. A surprisingly high proportion of respondents answered ‘Yes’ when asked if they had experienced auditory-evoked flashes while resting in the dark (39%), and these also rated the videos significantly higher [$F(2,1055) = 36.02$, $\eta^2_{part} = 0.064$]. Finally, significantly higher ratings were found with respondents who were 'Not Sure' they had tinnitus [$F(2,1055) = 5.52$, $p<0.004$, $\eta^2_{part} = 0.01$], while the proportion of 'Yes' respondents was unusually high (30%).

Motion energy analysis

Our *a priori* hypothesis was that two factors might influence ratings: firstly meaningful events depicted in the videos might be predictive of associated sounds (e.g. a screaming person's face, or collisions such as a hammer hitting a nail); secondly auditory sensations might be independently evoked by the amount of raw movement or just transient variations of light over time, or 'motion energy' (ME), regardless of the meaningful content of the video. A critical prediction from our hypothesis is that these two factors should each affect performance independently, as they relate to more cognitive versus more perceptual processes respectively. We tested this using principle components analysis of ratings averaged for each video. The first and second principle components (PCs) explained 61% and 6.9% of the data respectively. Further components made only a minimal further contribution (3%, 2.5%, 2%, etc.). Examining the distribution of videos over the first two PCs confirms our dual-factor hypothesis (Figure 4). Videos with higher coefficients for first PC tend to be meaningful and strongly predictive of real-world sounds, depicting rapid collisions, slow-motion bouncing, footsteps, and a person screaming. This latter video has very little motion in it. Videos high on the second PC included high contrast abstract moving patterns, twinkling lights, and flashing neon shop signs. Videos between these extremes tended to depict real-world scenes in rapid motion, which might also be associated with sounds, such as dancing, police car lights, and fireworks.

To verify that this second PC relates to the raw low-level motion, we correlated the PC coefficients of each video with motion energy quantified using a simple computational model of low-level motion processing (Adelson & Bergen, 1985) implemented using freely-available Matlab code (Mather, 2013). This model approximates the spatiotemporal filtering properties of the receptive fields of cells in early visual cortex, which respond to patterns on the retina moving in specific directions. One unusually high scoring video (z-score 3.68) was excluded from further analysis. As we predicted, motion energy estimates correlated significantly only with coefficients from PC2 [$r(21) = 0.63$, $p< 0.001$], not PC1 [$r(21) = -0.26$, $p< 0.224$].

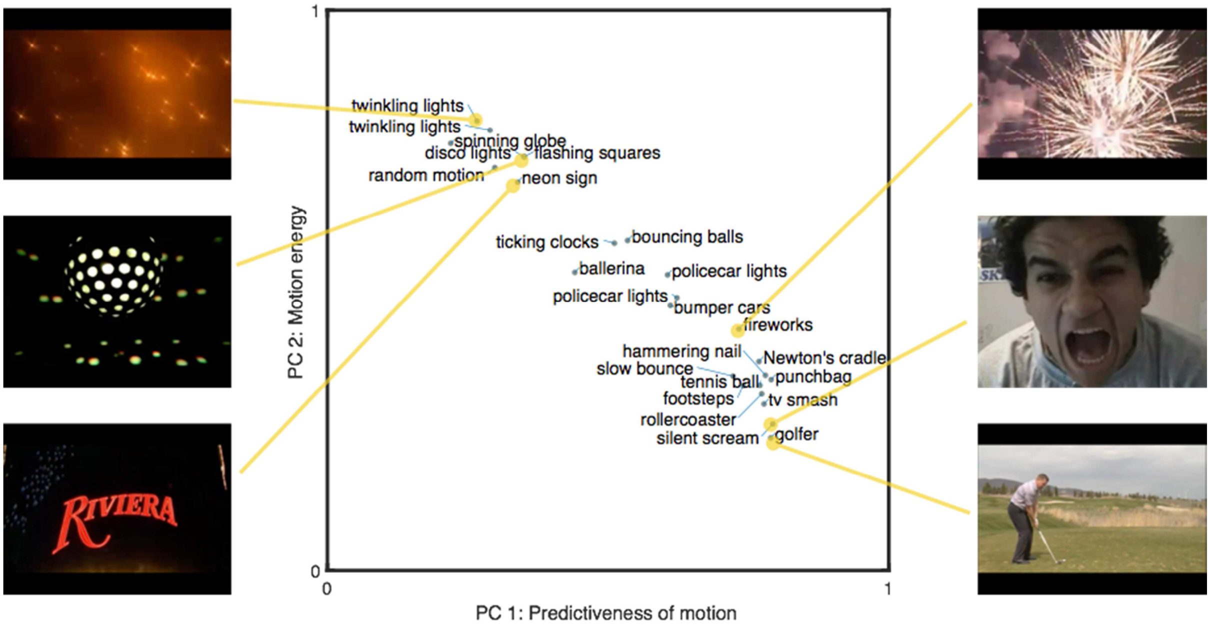


Figure 4 Coefficients for different videos on two principle components, after Varimax rotation, with stills from selected videos, and proposed interpretations of these dimensions shown as axis labels.

We next tested whether people who are sensitive to vEAR might rate videos higher if they have greater ME. We first median-split the stimuli into two sets, low and high ME, and compared mean ratings for each set grouping participants by their response to the question about previous awareness of vEAR. ANOVA results showed a significant interaction between previous awareness and ME [$F(1,4126) = 72.96$, $p<0.0001$, $\eta^2_{part} = 0.017$] (Figure 5, left bar chart). Both main effects were also significant [Previous awareness: $F(1,4126) = 722.40$, $p<0.0001$, $\eta^2_{part} = 0.149$, ME: $F(1,4126) = 212.70$, $p=0.0001$, $\eta^2_{part} < 0.049$]. We next used our PCA results to predict ratings for each video as a function of ME and participant group. We reconstructed ratings for each video based on PC2 and higher components, but excluding PC1 (i.e. Reconstructed ratings = PCscores(2^{nd}…last) x Eigenvectors(2^{nd}…last)^T), and split these data by previous awareness group, prior to averaging over participants. Reconstructed ratings averaged for each stimulus correlated significantly and positively with ME only for the ‘Yes’ awareness group [right scatterplot, $r(21) = 0.57$, $p< 0.005$]. Correlations with the other two groups were weakly negative [Not Sure: $r(21) = -.43$, $p=.04$, No: $r(21) = -.41$, $p=.05$]. Fisher’s Z tests showed that the correlation in the 'Yes' awareness group was significantly higher than for the other two groups ['Yes' vs 'No': Z=3.4, p = 0.00031; 'Yes' vs 'Not sure': Z = 3.5, p = 0.00023], which did not differ significantly from each other [Z=0.08]. The negative association found for the non-'Yes' respondents might relate to the choice of videos, which tended to have either high predictiveness but little movement, or more abstract movement. No significant correlations with ME were observed when ratings were reconstructed excluding PC2 but including PC1 instead (centre scatterplot). A similar pattern was observed when grouping respondents by their answer to the question about whether they consider themselves to be a synaesthete. 'Yes' respondents showed a significant positive correlation with ME [r(21) = 0.57, p= 0.0047], while for 'No' respondents the correlation was significantly negative [r(21) = -0.57, p= 0.0047].

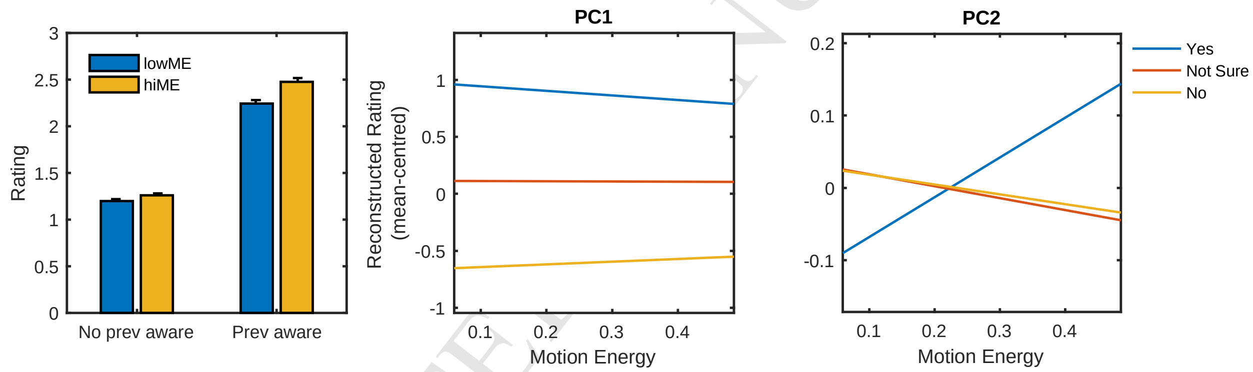


Figure 5 Sensitivity to motion energy in videos, for respondents differing in previous awareness of visually-evoked sounds. Left graph; average ratings for videos with low and high motion energy (different colours), split by previous awareness of vEAR. Centre and right graphs: scatterplots showing associations of motion energy with average ratings for each stimulus, reconstructed from either of two principle components, mean-centred, and split by previous awareness of vEAR (different colours).

Association of motion energy sensitivity with survey responses

While the previous analyses established that subjective reports of previous awareness of vEAR are associated with more positive survey responses, and also greater sensitivity to motion energy in the videos, a final analysis examined whether greater sensitivity to motion energy is itself associated with more positive survey responses. By replacing the 'previous awareness' question with an objective measure of stimulus motion, this analysis also avoids a potential source of response bias, whereby some participants might have tended answer more positively to the 'previous awareness' question and the other survey questions. We first examined the distribution of Pearson's r correlation coefficients between video ratings and motion energy for each respondent (Mean -.13, SD 0.26). The distributions differed significantly depending on responses to the 'previous awareness' question ['Yes': Mean -0.083 (SD 0.28); 'No': Mean -0.18 (SD 0.25), K-S test = 0.11, $p<0.0001$]. 38% of 'Yes' participants had positive r values, versus 25% of other participants, and 27% overall.

We then classified respondents into three groups depending on whether their r values were in the lowest quartile of the distribution of correlation coefficients (r values from -0.75 to -0.32), the top quartile (r values from .05 to .80), or from the central half of the distribution. We hypothesised that those in the top quartile would show similar patterns of association with survey responses to those reporting previous awareness of vEAR, if at least some the latter respondents genuinely experienced vEAR evoked by low-level visual motion energy.

Our results support this hypothesis, showing that sensitivity to motion energy had a significant effect on survey responses. Participants from the top quartile made slightly higher video ratings [bottom quartile: Mean 1.3, SE 0.033; middle: 1.48, SE 0.026; top: 1.63, SE 0.37; $F(2,4330) = 22.19$, $p<0.00001$, $\eta^2_{part} = 0.01$], and tended to report slightly more frequent musical imagery [$F(2,1067) = 3.09$, $p<.05$, $\eta^2_{part} = 0.006$]. They also tended to report more frequently auditory-evoked phosphenes [$\chi^2(4) = 15.24$, p=.004, V = .07], and tinnitus-like phenomena [$\chi^2(4) = 9.48$, p=.05, V = .05]. These effects are small, given that our quartile-based selection included respondents with weak motion energy correlations, and combined respondents who did and did not report previous awareness of vEAR. However these results establish that sensitivity to ME is associated with vEAR and other traits, independently of possible biases that might affect responses to the 'previous awareness' question.

Discussion

Since the initial small-scale report describing the ability to hear visual motion a decade ago (Saenz & Koch, 2008), this is the first large-scale study of this phenomenon to systematically examine both the visual factors that best evoke auditory sensations, and the individual traits that are associated with susceptibility. Our new results inform current debates about the mechanisms underlying synaesthesia-like abilities such as vEAR, and raise new hypotheses for future neuroscience research. They also suggest that vEAR might be surprisingly prevalent.

On the internet it is easy to find several large collections of silent movies which have been specifically selected for having the property of evoking illusory sounds (e.g. search for '*gifs you can hear*'). One site at www.reddit.com/r/noisygifs/ has over 54,000 subscribers. The existence of such sites and their popularity suggests that awareness of visually-evoked sounds among the general public is growing, despite the lack of scientific recognition or research into it. Our findings are consistent with this informal observation. In our main survey, 21% of our 4128 completing respondents, and 16% of 26118 non-completing, responded 'Yes' when asked about previous experience of vEAR. These respondents were self-selected, having been recruited via a news article reporting our prior research on vEAR. However, the proportion of 'Yes' respondents was even higher (31%) in a smaller group of naïve participants (N=126) recruited from a paid panel. As a further point of comparison from our previous study using a small randomly-selected naïve sample (Fassnidge et al., 2017), 22% confirmed that they had noticed sounds accompanying flash stimuli presented in the lab.

The validity might be questioned of classifying respondents on the basis of their answer to a single subjective question about previous awareness of vEAR. As a check of consistency we asked a similar question at the end of the survey about experiences of hearing flashes and movements in everyday life, and we found similar results in all of our analyses when including only the participants who made the same response to both questions. However to address this concern about validity, we have adopted an alternative method used in previous study by Rothen et al (2017), in which participants were identified as experiencing vEAR if they confirmed that they experienced auditory sensations (‘Yes’ or ‘No’), in at least half of 12 videos shown. This resulted in an estimated prevalence of only 4.2% out of 221 naïve participants. Our own attempt to use a similar ratings-based criterion found that 22% of our naïve respondents gave ratings greater than or equal to 3 (out of maximum rating of 5) in at least half of the videos presented. Two factors might explain why our estimate is higher. Firstly, Rothen et al used binary ‘Yes/No’ options for responding to videos, which might have induced a conservative bias, while we used a 6-point rating scale. Secondly Rothen et al included questions which compared the intensity of visually-evoked sounds to examples of real sounds such as ‘whisper’, ‘street noise’, or ‘machinery’. This might have directed attention away from internally-generated auditory sensations. In contrast, we emphasised this internal source in our introductory text, which could partially account for our higher estimates. However, even Rothen et al’s conservative estimate suggests that vEAR may have relatively high prevalence in comparison to a previous estimates of 4.4% for having any other kind of synaesthesia (Simner et al., 2006).

It might also be questioned whether such answers truly refer to a synaesthesia-like visually-induced auditory sensation rather than a high-level form of visually-associated auditory imagery or crossmodal correspondence. In particular, a tendency to answer ‘Yes’ might have been inflated given that in the absence of a consensus on the phenomenology of vEAR, the definition we provided prior to the question included references to imagination, ringing in the ears, and other phenomenological features. We preferred to define vEAR broadly in this study, because there has not yet been any specific consensus on the phenomenology, and different individuals may interpret the sensations differently. In addition it may be over-restrictive to apply standardised criteria of the kind used to diagnose canonical forms of synaesthesia (Ward, 2013). For example in grapheme-colour synaesthesia each specific letter of the alphabet might automatically and consistently induce a different specific colour, while the mapping is idiosyncratic to each individual. However in our informal interviews with individuals describing their experience of vEAR, different visual images or movements may evoke similar generic ‘white noise’ or ‘whooshing’ sounds rather than distinct and elaborate auditory imagery. Similar accounts were published in Rothen et al (2013) and Saenz & Koch (2008). Such a many-to-few mapping also seems qualitatively distinct from the kinds of specific cross-modal correspondences that are often reported between congruent sensory dimensions, for example associating auditory pitch with elevation and size, or visual shape with phonemic sounds (Spence, 2011). One empirical point in favour of vEAR being a genuine form of synaesthesia is that reports of previous awareness were associated with reports of other forms of synaesthesia (Barnett et al., 2008; Rothen et al., 2013; Sagiv et al., 2006), although again we did not formally assess these synaesthetic tendencies. Future research eliciting more detailed subjective reports of the specific quality of auditory sensation and other synaesthesias, may help to establish with greater certainty whether vEAR is really a kind of synaesthesia, or a vivid form of cross-modal correspondence, or whether the diagnostic criteria may need to be extended to include cases where similar generic sensory phenomena are evoked by different inducers.

A further concern about the validity of our measures is that responses might have been biased by 'yea-saying' behaviour, whereby some respondents might tend to respond more positively to the question about 'previous awareness' (Simner et al., 2006). This might arguably account for the observed positive association of such responses with traits such as musical imagery, auditory-evoked phosphenes and other synaesthetic phenomena. Not all responses are consistent with such a bias, for example previous awareness was associated more with 'Not Sure' than 'Yes' responses to the tinnitus question. However this concern may be addressed using an analysis of the correlation of video ratings with an objective measure of the motion energy in the videos. Specifically, we found that the video ratings of 'Yes' respondents to the 'previous awareness' question were more positively correlated with the amount of pure motion energy (ME) present even in abstract meaningless videos, such as swirling or patterns that were not predictive of sounds (Figure 5). Conversely, we also found that higher correlations of ratings with ME predicted higher video ratings across all participants, higher frequency of reporting previous awareness of vEAR and synaesthesia more generally, and reports of other traits such as musical imagery, tinnitus-like sounds and auditory-evoked phosphenes. Yea-saying would be unlikely to bias responses specifically towards videos with higher ME, and therefore cannot fully account for these correlations. Instead it seems more likely that survey responses at least partially reflect a common factor related to sensitivity to visual motion energy. This analysis does not provide a criterion-free estimate of prevalence, however approximately 13% more ‘previously aware’ respondents had positive correlations with ME compared to other respondents (38% versus 25%). These results suggest that at least some of the ‘previous awareness’ respondents were genuinely reporting sensations evoked by low-level visual motion, rather than merely reflecting a high-level association of meaningful visual events and their predicted sounds, or a bias towards responding positively to survey questions. In addition they provide further evidence of specific associations between this low-level form of vEAR and our other trait measures, avoiding potential sources of response bias.

This association between survey responses and motion energy might inform the debate (mentioned in the introduction) about the neuro-architectural underpinnings of such experiences (Hubbard & Ramachandran, 2005). There may be a relatively direct effect of low-level visual stimulation on auditory processing (Kayser, Petkov, & Logothetis, 2008; Schroeder et al., 2008), which need not depend on access to higher-level semantic representations (Mattingley et al., 2001). Further evidence supports this low-level route: simple motion energy inducers evoke simple concurrents (Rothen et al, 2017; e.g. generic ‘whooshing’ noises) while early visual-evoked potentials may also be enhanced in people identified as experiencing vEAR (Rothen et al., 2017); furthermore, we previously found that watching simple abstract visual motion may involuntarily disrupt detection of real sounds (Fassnidge et al., 2017). As an independent factor, higher ratings were also given to videos depicting events that are naturally associated with sounds, such as collisions, in both ‘Yes’ and ‘No’ respondents. It is unclear how much this strong effect reflects cognitive imagery versus actual perceptions of sounds, however it seems consistent with the co-existence of a high-level semantically-mediated route to vEAR (Mattingley et al., 2001; Myles et al., 2003; Smilek et al., 2001).

Our results also inform a second ongoing debate about the neurophysiological mechanisms underlying synaesthesia, in particular whether such phenomena depend on unusual connectivity between specific brain areas (Bargary & Mitchell, 2008; Ramachandran & Hubbard, 2001; Rouw & Scholte, 2007), versus modulation of inhibition or excitation within the context of normal connectivity (Brang, Williams, & Ramachandran, 2012; Cohen Kadosh, Henik, Catena, Walsh, & Fuentes, 2009;

Grossenbacher & Lovelace, 2001; Hubbard, Brang, & Ramachandran, 2011). It has been argued that unusual connectivity can result from insufficient pruning of specific cortical interconnections (Baron-Cohen, 1996) during the development of individuals representing relatively rare genotypes (Tomson et al., 2011), leading to unusual patterns of neural connectivity between specific brain areas. This framework may explain some canonical forms of synaesthesia between usually unrelated stimulus dimensions such as graphemes and colour (Rouw & Scholte, 2007), as well as their rarity and their idiosyncratic variety. However, an account in which rarely-occurring and idiosyncratic mutations contribute to unusual patterns of connectivity does not easily explain the apparently higher prevalence of vEAR relative to other synaesthesias, nor its apparently broad association with other traits and forms of synaesthesia.

Richer connectivity between visual and auditory areas may not be so rare, however, because auditory and visual events are much more highly correlated in nature (e.g. whenever two objects collide or a person speaks) than the dimensions normally associated with synaesthesia such as colour and visual forms (Fassnidge et al., 2017). Hebbian learning of these associations during normal development might then reinforce functional audiovisual connectivity and protect against neural pruning (Baron-Cohen, 1996). The survival of connections between auditory and visual modalities through to adulthood might explain not just synaesthesia-like perceptions in a subset of individuals, but the existence of other phenomena such as auditory-evoked visual phosphenes, as well as other highly commonplace associations between vision and audition. For example, we like to listen to music synchronised with flashing lights or dance, or violins while watching a conductor's gestures, and incidental sound effects accompanying action in movies, such as the comic 'boing' when a cartoon character slips on a banana skin. These stimuli might reinforce each other via such audiovisual connections.

In contrast to the above argument, abundance of audiovisual connectivity may not be sufficient to explain our findings that subjective reports of vEAR, and sensitivity to motion energy, are associated with reports of unimodal phenomena such as tinnitus-like internal sounds and musical imagery, as well as reports of other forms of synaesthesia. Such reports should be treated with caution as we did not undertake a detailed assessment of these traits using standard measures in the interests of not overburdening our voluntary respondents. However our finding that they are more slightly prevalent among respondents who are sensitive to motion energy reinforces their validity as a correlate of vEAR. These generalised phenomena might additionally depend on systemic variations in cortical excitability or disinhibition of sensory brain areas or their interconnections (Brang et al., 2012; Grossenbacher & Lovelace, 2001; Neufeld et al., 2012). Other synaesthetic and cross-modal phenomena have been explained by disinhibited feedback from multisensory areas such as inferior parietal cortex or superior temporal sulcus to unimodal areas (Grossenbacher & Lovelace, 2001; Neufeld et al., 2012). Such disinhibited feedback might explain the audiovisual phenomenology of vEAR and its association with the reverse direction of auditory-evoked phosphenes, but is less consistent with our finding of associations with unimodal phenomena such as tinnitus-like internal sounds and musical imagery. These associations suggests that aside from disinhibition of specific feedback, the local response of sensory areas may also be disinhibited or rendered more excitable. This might render them more responsive to inputs via feedback, horizontal or bottom-up connections, and possibly to spontaneous neural firing.

Our questions about musical imagery, tinnitus-like phenomena and auditory-evoked phosphenes were included to test this excitability hypothesis. We did not intend to establish a formal restrictive diagnosis of these phenomena, and this limits how confidently we can interpret responses to these questions as reflecting genuine experiences. Instead, our aim was to capture a wide range of experiences relating to the sensory phenomena of interest, which we predicted would covary with vEAR. Our choice of questions was motivated by past evidence that involuntary musical imagery may arise from increased spontaneous activity in auditory areas especially following hearing loss (Kumar et al., 2014), while some varieties of tinnitus may also reflect disinhibition of auditory areas (Kaltenbach, 2011). Auditory-evoked phosphenes, sometimes referred to as ‘auditory-visual synaesthesia’, have previously been reported mostly in patients with pathologically reduced visual input (Jacobs et al., 1981; Lessell & Cohen, 1979), and the hypnagogic variety might occur due to increased excitability of visual cortex under dark adaptation, as evidenced by lower thresholds for inducing phosphenes by magnetic stimulation of occipital lobe (Boroojerdi et al., 2000).

vEAR awareness was positively associated with frequency of reporting tinnitus-like phenomena, however video ratings were actually highest in respondents who were ‘not sure’ about having tinnitus. This might be explained if faint visually-evoked sounds were drowned out by loud tinnitus and more noticeable in borderline conditions, or if there is an experience associated with heightened auditory excitability which is phenomenologically distinct from the symptoms of true tinnitus. The interpretation of this result however remains uncertain because we did not undertake a formal clinical assessment of tinnitus. Hypnagogic auditory-evoked visual phosphenes were also reliably associated with higher ratings, prior awareness of vEAR, and sensitivity to motion energy. Spontaneous hypnagogic imagery is not uncommon, and past research has noted an association of this with tendencies for synaesthesia (Terhune, 2009), however our survey provides the first assessment of hypnagogic phosphenes evoked specifically by auditory stimulation. A surprisingly high proportion of respondents (39%) reported this phenomenon, although an confident assessment of prevalence would require a more detailed investigation.

The association of these traits with vEAR is consistent with past suggestions that excitability might play a common role in modulating the response to cross-modal signals (Bolognini et al., 2010; Schroeder et al., 2008; Terhune et al., 2011). In support it has been found that experimental manipulation of cortical excitability via brain stimulation may modulate colour-grapheme (Terhune et al., 2011) and mirror-touch synaesthesia (Bolognini, Miniussi, Gallo, & Vallar, 2013), as well as effects of auditory stimuli on visual phosphenes (Bolognini et al., 2010). Further support for an excitability or disinhibition explanation, specifically for vEAR, comes from Rothen et al’s (2017) recent study showing enhanced visual evoked potentials (N2 and earlier) in participants reporting visually-evoked sounds. One candidate mechanism for this is suggested by pharmacological studies showing that synaesthesia and visual phosphene thresholds, as well as migraine aura and hallucinations all depend on the action of serotonin (5-HT) primarily via S2a receptors (Aghajanian & Marek, 1999; Brang & Ramachandran, 2008; Brogaard, 2013; Hamel, 2007; Luke & Terhune, 2013; Oliveri, 2003) which may have complex effects on cortical excitability by acting on glutamate- and GABA-mediated transmission (Ciranna, 2006).

In summary, our combined analysis of stimulus and trait factors associated with vEAR suggest that auditory sensations might be evoked by low-level abstract visual motion energy without necessarily requiring higher-level interpretative processes; furthermore, individual differences in levels of disinhibition or excitability within a network of audiovisual brain areas may result in a variety of associated conscious experiences, including sounds evoked by flashes, flashes evoked by sounds, and sounds resembling tinnitus or music. This explanatory hypothesis may now be tested by correlating behavioural measures of vEAR and these other related phenomena with measures of brain connectivity and excitability, or by using vEAR as a dependent measure of the effects of experimental pharmacology or brain stimulation. Given the prevalence of visual-evoked auditory sensations and our new ability to quantify and correlate them, a potentially broad class of related subjective audiovisual phenomena have now become more accessible to scientific study.

Acknowledgements
This research was supported by British Academy/Leverhulme grant SG151380. We thank Danai Dima, Alberta Ipser, Marinella Cappelletti, James Hampton, Jamie Ward and two anonymous reviewers for helpful comments.

References

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America. A, Optics and Image Science*, *2*(2), 284–99.

Aghajanian, G. K., & Marek, G. J. (1999). Serotonin and hallucinogens. *Neuropsychopharmacology*, *21*, 16S–23S. http://doi.org/10.1016/S0893-133X(98)00135-3

Banissy, M. J., Cassell, J. E., Fitzpatrick, S., Ward, J., Walsh, V. X., & Muggleton, N. G. (2012). Increased positive and disorganised schizotypy in synaesthetes who experience colour from letters and tones. *Cortex*, *48*(8), 1085–1087. http://doi.org/10.1016/j.cortex.2011.06.009

Banissy, M. J., Holle, H., Cassell, J., Annett, L., Tsakanikos, E., Walsh, V., … Ward, J. (2013). Personality traits in people with synaesthesia : Do synaesthetes have an atypical personality profile ?, *54*, 828–831.

Banissy, M. J., Walsh, V., & Ward, J. (2009). Enhanced sensory perception in synaesthesia. *Experimental Brain Research*, *196*(4), 565–571. http://doi.org/10.1007/s00221-009-1888-0

Bargary, G., & Mitchell, K. J. (2008). Synaesthesia and cortical connectivity. *Trends in Neurosciences*. http://doi.org/10.1016/j.tins.2008.03.007

Barnett, K. J., Finucane, C., Asher, J. E., Bargary, G., Corvin, A. P., Newell, F. N., & Mitchell, K. J. (2008). Familial patterns and the origins of individual differences in synaesthesia. *Cognition*, *106*(2), 871–893. http://doi.org/10.1016/j.cognition.2007.05.003

Baron-Cohen, S. (1996). Is there a normal phase of synaesthesia in development. *Psyche*, *2*(27), 223–228.

Baron-Cohen, S., Johnson, D., Asher, J., Wheelwright, S., Fisher, S. E., Gregersen, P. K., & Allison, C. (2013). Is synaesthesia more common in autism? *Molecular Autism*, *4*(1), 40. http://doi.org/10.1186/2040-2392-4-40

Bolognini, N., Miniussi, C., Gallo, S., & Vallar, G. (2013). Induction of mirror-touch synaesthesia by increasing somatosensory cortical excitability. *Current Biology*, *23*(10), R436–R437. http://doi.org/10.1016/j.cub.2013.03.036

Bolognini, N., Senna, I., Maravita, A., Pascual-Leone, A., & Merabet, L. B. (2010). Auditory enhancement of visual phosphene perception: the effect of temporal and spatial factors and of stimulus intensity. *Neuroscience Letters*, *477*(3), 109–14. http://doi.org/10.1016/j.neulet.2010.04.044

Bor, D., Rothen, N., Schwartzman, D. J., Clayton, S., & Seth, A. K. (2014). Adults can be trained to acquire synesthetic experiences. *Scientific Reports*, *4*, 7089. http://doi.org/10.1038/srep07089

Boroojerdi, B., Bushara, K. O., Corwell, B., Immisch, I., Battaglia, F., Muellbacher, W., & Cohen, L. G. (2000). Enhanced excitability of the human visual cortex induced by short-term light deprivation. *Cereb Cortex*, *10*(5), 529–534. http://doi.org/10.1093/cercor/10.5.529

Brang, D., & Ramachandran, V. S. (2008). Psychopharmacology of synesthesia; the role of serotonin S2a receptor activation. *Medical Hypotheses*, *70*(4), 903–904. http://doi.org/10.1016/j.mehy.2007.09.007

Brang, D., Williams, L. E., & Ramachandran, V. S. (2012). Grapheme-color synesthetes show enhanced crossmodal processing between auditory and visual modalities. *Cortex*, *48*(5), 630–7. http://doi.org/10.1016/j.cortex.2011.06.008

Brogaard, B. (2013). Serotonergic Hyperactivity as a Potential Factor in Developmental, Acquired and Drug-Induced Synesthesia. *Frontiers in Human Neuroscience*, *7*(October), 1–13. http://doi.org/10.3389/fnhum.2013.00657

Carmichael, D. A., & Simner, J. (2013). The immune hypothesis of synesthesia. *Frontiers in Human Neuroscience*, *7*(September), 1–3. http://doi.org/10.3389/fnhum.2013.00563

Ciranna, L. (2006). Serotonin as a Modulator of Glutamate- and GABA-Mediated Neurotransmission: Implications in Physiological Functions and in Pathology. *Current Neuropharmacology*, *4*(2), 101–114. http://doi.org/10.2174/157015906776359540

Cohen Kadosh, R., Henik, A., Catena, A., Walsh, V., & Fuentes, L. J. (2009). Induced cross-modal synaesthetic experience without abnormal neuronal connections. *Psychological Science*, *20*(2), 258–65. http://doi.org/10.1111/j.1467-9280.2009.02286.x

Devlin, H. (2017, January 17). Listen with your eyes: one in five of us may “hear” flashes of light. Retrieved September 15, 2017, from https://www.theguardian.com/science/2017/jan/17/listen-with-your-eyes-one-in-five-of-us-may-hear-flashes-of-light-synaesthesia

Fassnidge, C., Cecconi Marcotti, C., & Freeman, E. (2017). A deafening flash! Visual interference of auditory signal detection. *Consciousness and Cognition*, *49*, 15–24. http://doi.org/10.1016/j.concog.2016.12.009

Grossenbacher, P. G., & Lovelace, C. T. (2001). Mechanisms of synesthesia: cognitive and physiological constraints. *Trends in Cognitive Sciences*, *5*(1), 36–41. http://doi.org/10.1016/S1364-6613(00)01571-0

Hamel, E. (2007). Serotonin and migraine: biology and clinical implications. *Cephalalgia*, *27*(11), 1293–1300.

Hubbard, E. M., Brang, D., & Ramachandran, V. S. (2011). The cross-activation theory at 10. *Journal of Neuropsychology*, *5*(2), 152–177. http://doi.org/10.1111/j.1748-6653.2011.02014.x

Hubbard, E. M., & Ramachandran, V. S. (2005). Neurocognitive mechanisms of synesthesia. *Neuron*, *48*(3), 509–520. http://doi.org/10.1016/j.neuron.2005.10.012

Jacobs, L., Karpik, A., Bozian, D., & Gøthgen, S. (1981). Auditory-visual synesthesia sound-induced photisms. *Archives of Neurology*, *38*(4), 211–216.

Kaltenbach, J. A. (2011). Tinnitus: Models and mechanisms. *Hearing Research*, *276*(1–2), 52–60. http://doi.org/10.1016/j.heares.2010.12.003

Kayser, C., Petkov, C. I., & Logothetis, N. K. (2008). Visual modulation of neurons in auditory cortex. *Cerebral Cortex*, *18*(7), 1560–1574. http://doi.org/10.1093/cercor/bhm187

Kumar, S., Sedley, W., Barnes, G. R., Teki, S., Friston, K. J., & Griffiths, T. D. (2014). A brain basis for musical hallucinations. *Cortex*, *52*, 86–97. http://doi.org/10.1016/j.cortex.2013.12.002

Lessell, S., & Cohen, M. M. (1979). Phosphenes induced by sound. *Neurology*, *29*(11), 1524.

Luke, D. P., & Terhune, D. B. (2013). The induction of synaesthesia with chemical agents: a systematic review. *Frontiers in Psychology*, *4*(October), 753. http://doi.org/10.3389/fpsyg.2013.00753

Mather, G. (2013). Matlab implementation of the Adelson-Bergen motion energy sensor. Retrieved September 15, 2017, from http://www.georgemather.com/Model.html

Mattingley, J. B., Rich, A. N., Yelland, G., & Bradshaw, J. L. (2001). Unconscious priming eliminates automatic binding of colour and alphanumeric form in synaesthesia. *Nature*, *410*, 580–582. http://doi.org/10.1038/35069062

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature*, *264*(5588), 746–748. http://doi.org/10.1038/264746a0

Myles, K. M., Dixon, M. J., Smilek, D., & Merikle, P. M. (2003). Seeing double: The role of meaning in alphanumeric-colour synaesthesia. *Brain and Cognition*, *53*(2), 342–345. http://doi.org/10.1016/S0278-2626(03)00139-8

Neufeld, J., Sinke, C., Zedler, M., Dillo, W., Emrich, H. M., Bleich, S., & Szycik, G. R. (2012). Disinhibited feedback as a cause of synesthesia: evidence from a functional connectivity study on auditory-visual synesthetes. *Neuropsychologia*, *50*(7), 1471–7. http://doi.org/10.1016/j.neuropsychologia.2012.02.032

Oliveri, M. (2003). Increased visual cortical excitability in ecstasy users: a transcranial magnetic stimulation study. *Journal of Neurology, Neurosurgery & Psychiatry*, *74*(8), 1136–1138. http://doi.org/10.1136/jnnp.74.8.1136

Ramachandran, V. S., & Hubbard, E. M. (2001). Psychophysical investigations into the neural basis of synaesthesia. *Proceedings of the Royal Society B: Biological Sciences*, *268*(1470), 979–983. http://doi.org/10.1098/rspb.2000.1576

Rothen, N., Bartl, G. J., Franklin, A., & Ward, J. (2017). Electrophysiological Correlates and Psychoacoustic Characteristics of Hearing-Motion Synaesthesia. *Neuropsychologia*, (in press). http://doi.org/https://doi.org/10.1016/j.neuropsychologia.2017.08.031

Rothen, N., Nikolić, D., Jürgens, U. M., Mroczko-Wasowicz, A., Cock, J., & Meier, B. (2013). Psychophysiological evidence for the genuineness of swimming-style colour synaesthesia. *Consciousness and Cognition*, *22*(1), 35–46. http://doi.org/10.1016/j.concog.2012.11.005

Rouw, R., & Scholte, H. S. (2007). Increased structural connectivity in grapheme-color synesthesia. *Nature Neuroscience*, *10*(6), 792–797. http://doi.org/10.1038/nn1906

Rouw, R., & Scholte, H. S. (2016). Personality and cognitive profiles of a general synesthetic trait. *Neuropsychologia*, *88*, 35–48. http://doi.org/10.1016/j.neuropsychologia.2016.01.006

Saenz, M., & Koch, C. (2008). The sound of change: visually-induced auditory synesthesia.

Current Biology, *18*(15), 650–651.

Sagiv, N., Simner, J., Collins, J., Butterworth, B., & Ward, J. (2006). What is the relationship between synaesthesia and visuo-spatial number forms? *Cognition*, *101*(1), 114–128. http://doi.org/10.1016/j.cognition.2005.09.004

Schroeder, C. E., Lakatos, P., Kajikawa, Y., Partan, S., & Puce, A. (2008). Neuronal oscillations and visual amplification of speech. *Trends in Cognitive Sciences*, *12*(3), 106–13. http://doi.org/10.1016/j.tics.2008.01.002

Simner, J., Mulvenna, C., Sagiv, N., Tsakanikos, E., Witherby, S. A., Fraser, C., … Ward, J. (2006). Synaesthesia: The prevalence of atypical cross-modal experiences. *Perception*, *35*(8), 1024–1033.

Smilek, D., Dixon, M. J., Cudahy, C., & Merikle, P. M. (2001). Synaesthetic photisms influence visual perception. *Journal of Cognitive Neuroscience*, *13*(7), 930–936. http://doi.org/10.1162/089892901753165845

Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, (January), 971–995. http://doi.org/10.3758/s13414-010-0073-7

Sumby, W. H., & Pollack, I. (1954). Visual Contribution to Speech Intelligibility in Noise. *Journal of the Acoustical Society of America*, *26*(2), 212–215. http://doi.org/10.1121/1.1907309

Terhune, D. B. (2009). The incidence and determinants of visual phenomenology during out-of-body experiences. *Cortex*, *45*(2), 236–242. http://doi.org/10.1016/j.cortex.2007.06.007

Terhune, D. B., Tai, S., Cowey, A., Popescu, T., & Cohen Kadosh, R. (2011). Enhanced cortical excitability in grapheme-color synesthesia and its modulation. *Current Biology*, *21*(23), 2006–9. http://doi.org/10.1016/j.cub.2011.10.032

Tomson, S. N., Avidan, N., Lee, K., Sarma, A. K., Tushe, R., Milewicz, D. M., … Eagleman, D. M. (2011). The genetics of colored sequence synesthesia: suggestive evidence of linkage to 16q and genetic heterogeneity for the condition. *Behav Brain Res*, *223*(1), 48–52. http://doi.org/10.1016/j.bbr.2011.03.071

Tsay, C.-J. (2013). Sight over sound in the judgment of music performance. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(36), 14580–5. http://doi.org/10.1073/pnas.1221454110

Ward, J. (2013). Synesthesia. *Annual Review of Psychology*, *64*(1), 49–75. http://doi.org/10.1146/annurev-psych-113011-143840

Ward, J., Hoadley, C., Hughes, J. E. A., Smith, P., Allison, C., Baron-Cohen, S., & Simner, J. (2017). Atypical sensory sensitivity as a shared feature between synaesthesia and autism. *Scientific Reports*, *7*(March), 41155. http://doi.org/10.1038/srep41155

Witthoft, N., Winawer, J., & Eagleman, D. M. (2015). Prevalence of Learned Grapheme-Color Pairings in a Large Online Sample of Synesthetes. *PLoS One*, *10*(3), e0118996. http://doi.org/10.1371/journal.pone.0118996